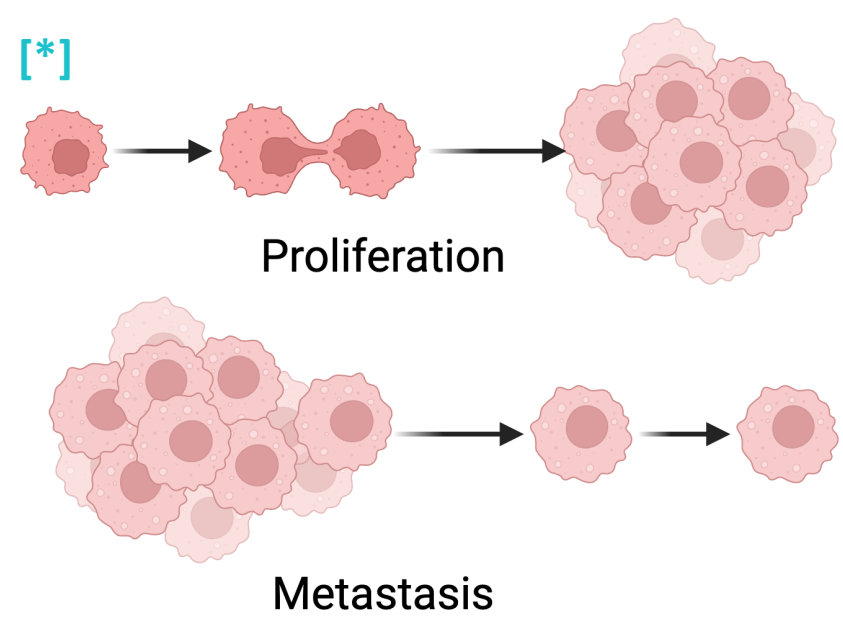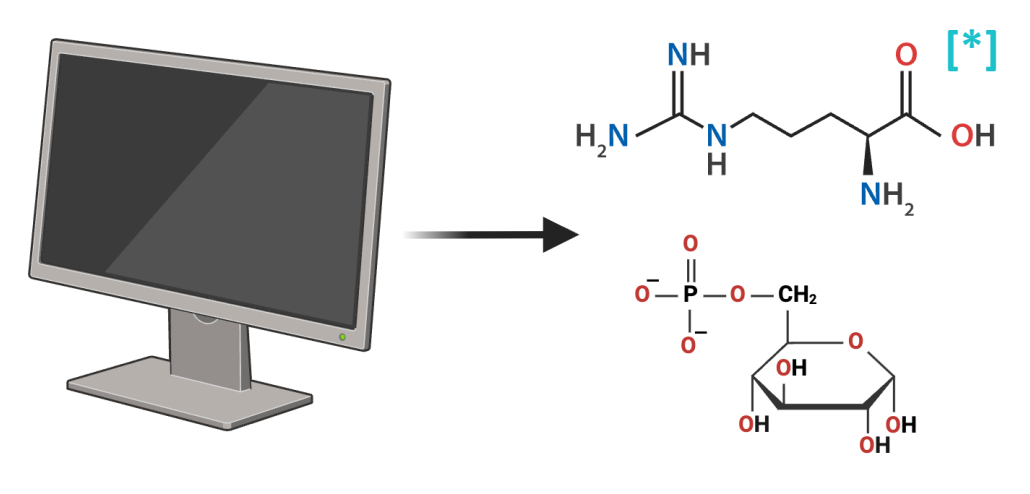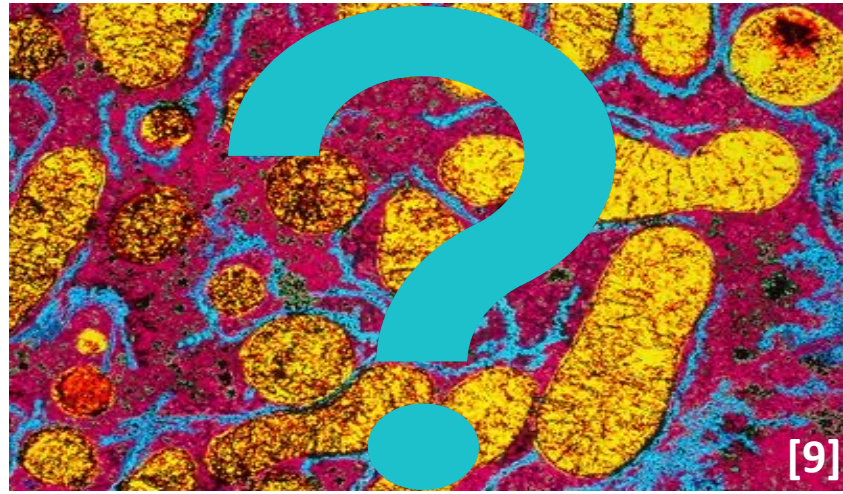# Overlooked Covariates in Metabolite Abundances: Systematically Quantifying the Information Overlap between Gene Expression and Metabolism across multiple cancer types

Introduction

[*]

Proliferation

Metastasis

[9]

Reprogrammed Metabolism

- **Essential for tumor proliferation and metastasis**[1]
- **Assists in immune evasion**[2]

[*]

Unable to effectively computationally predict or experimentally measure metabolism[3]

There has been no systematic quantification of the predictive power of gene expression, nor of the potential of other metadata as metabolic regulators

Methodology

Sourcing relevant and applicable datasets

Cancer Cell Line Encyclopedia (CCLE)[4]	Cancer Atlas of Metabolic Profiles (CAMP)[5]

Exploring the effect of covariates on metabolomic variance

 [6] [7]

Developing standardized preprocessing and curation

Metabolism	RNA

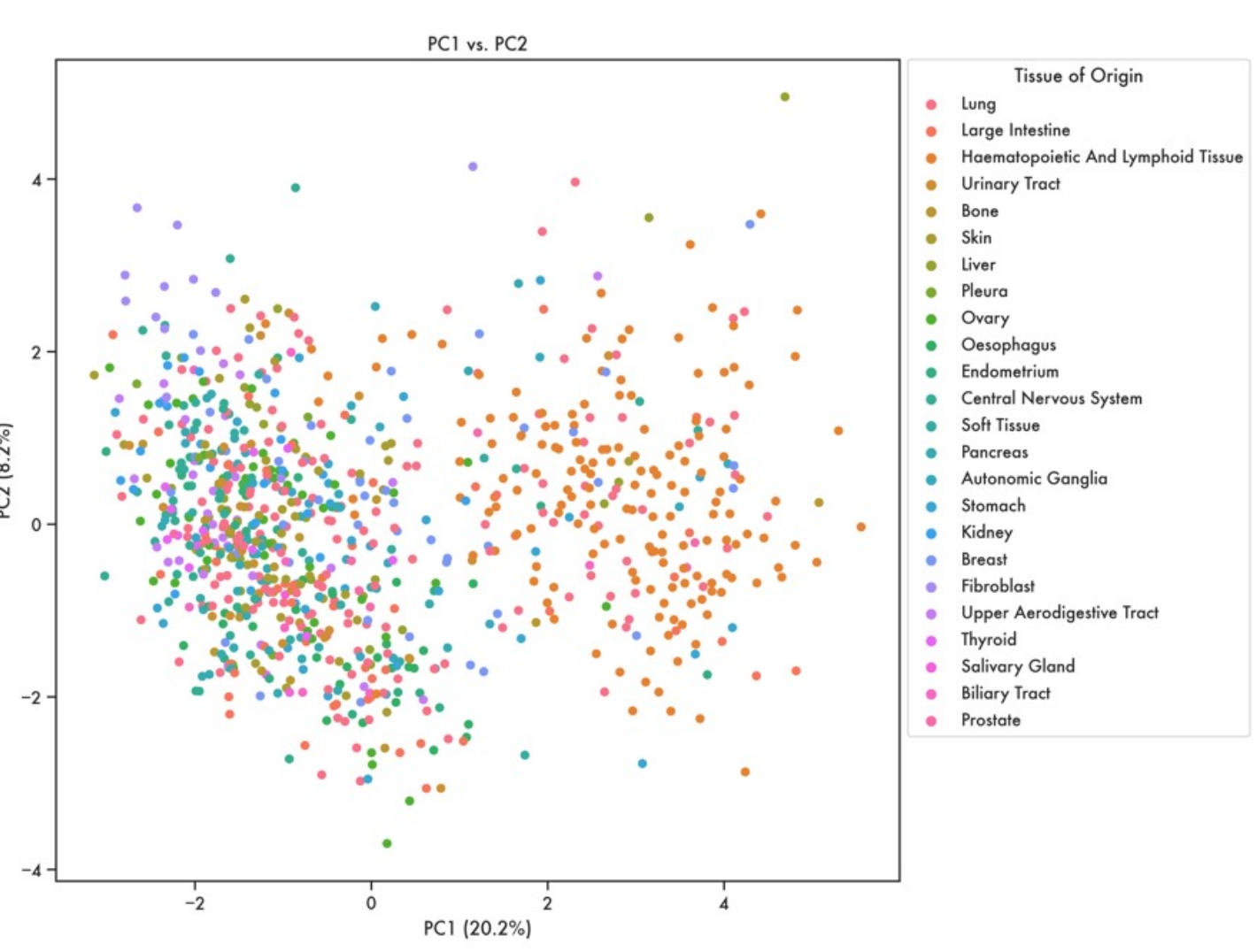### Identifying information overlap of gene expression and covariates

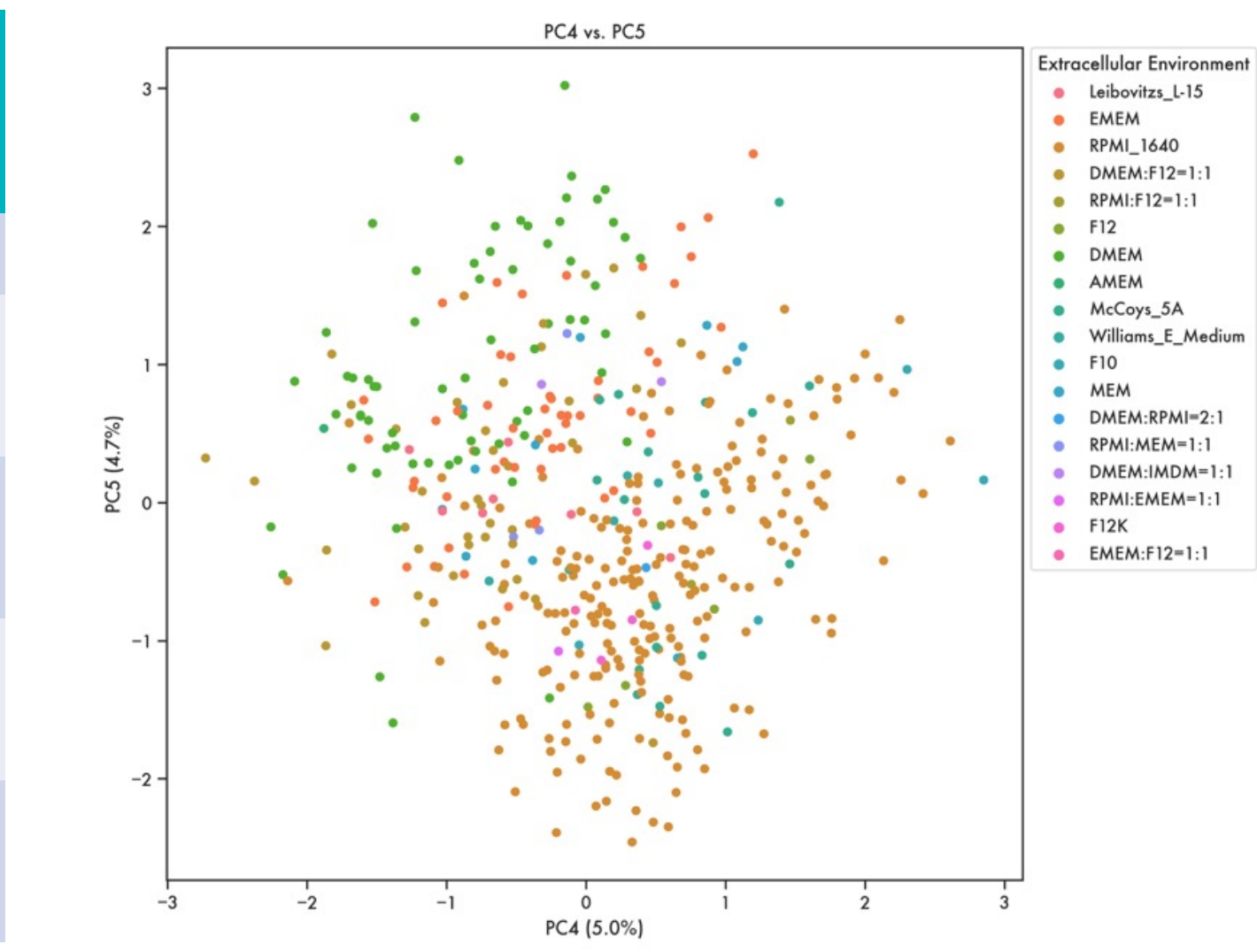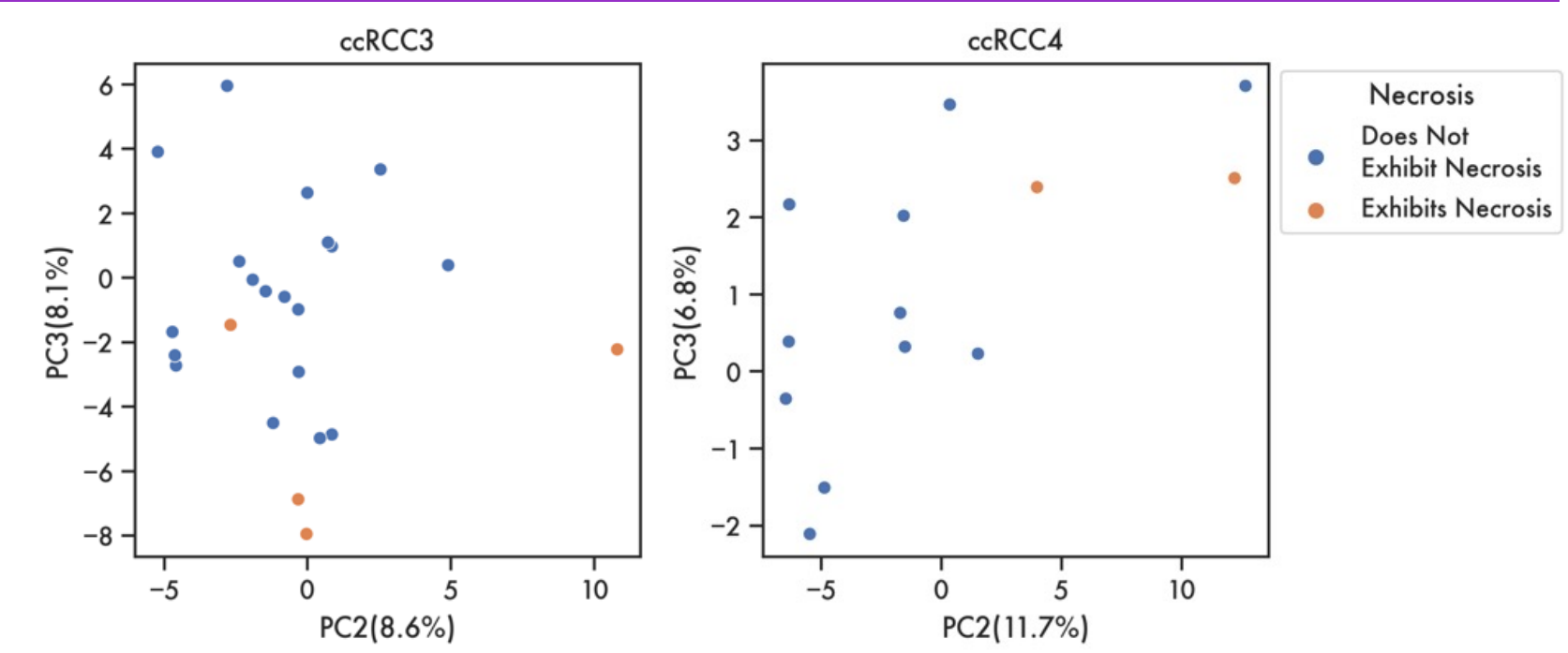| Distance Correlation (DC)[8]	Partial Distance Correlation (pDC)[8]
Novel non-linear metric of information overlap	Expresses proportion of DC related to covariates

Results

PC1 vs. PC2

PC	Associated Covariate/s	% of Variance
1	Tissue of Origin	20.2%
2	Growth Rate, Mutation Rate	8.2%
3	Growth Rate, Mutation Rate	7.1%
4	Extracellular Environment	5.0%
5	Extracellular Environment	4.7%

PC4 vs. PC5

Cancer Cell Line Encyclopedia

Case 2 Distance Correlation Mean Difference Across Normalizations
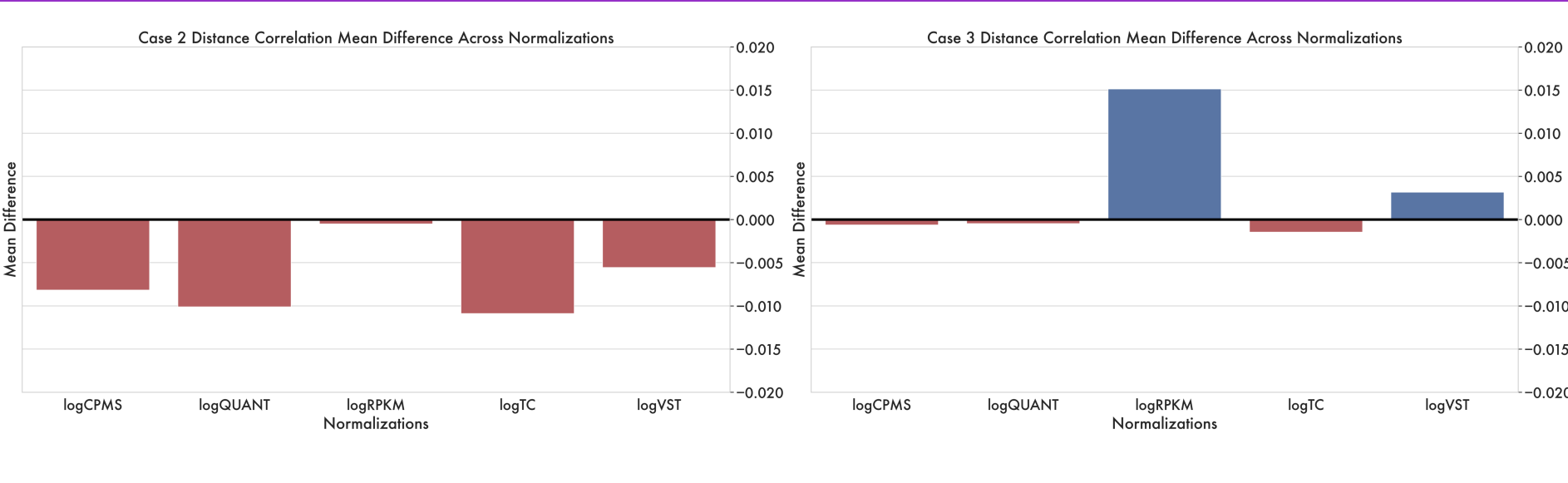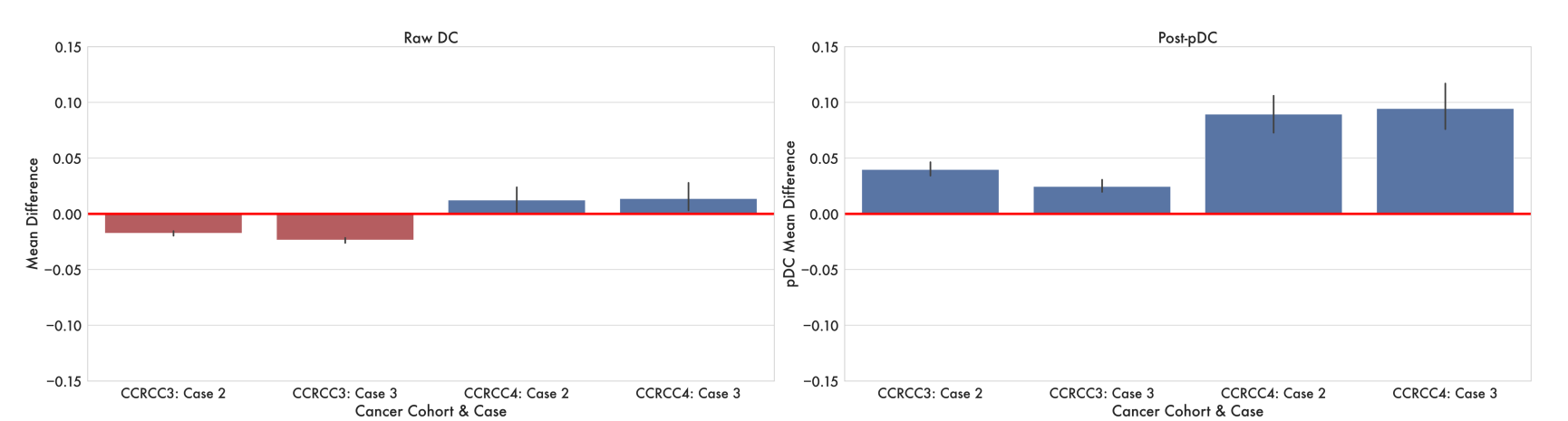

Case 3 Distance Correlation Mean Difference Across Normalizations

Sampling testing reveals more metabolic information is being associated with random genes than metabolic genes

Case 2 (Post-Removal)

Case 3 (Post-Removal)
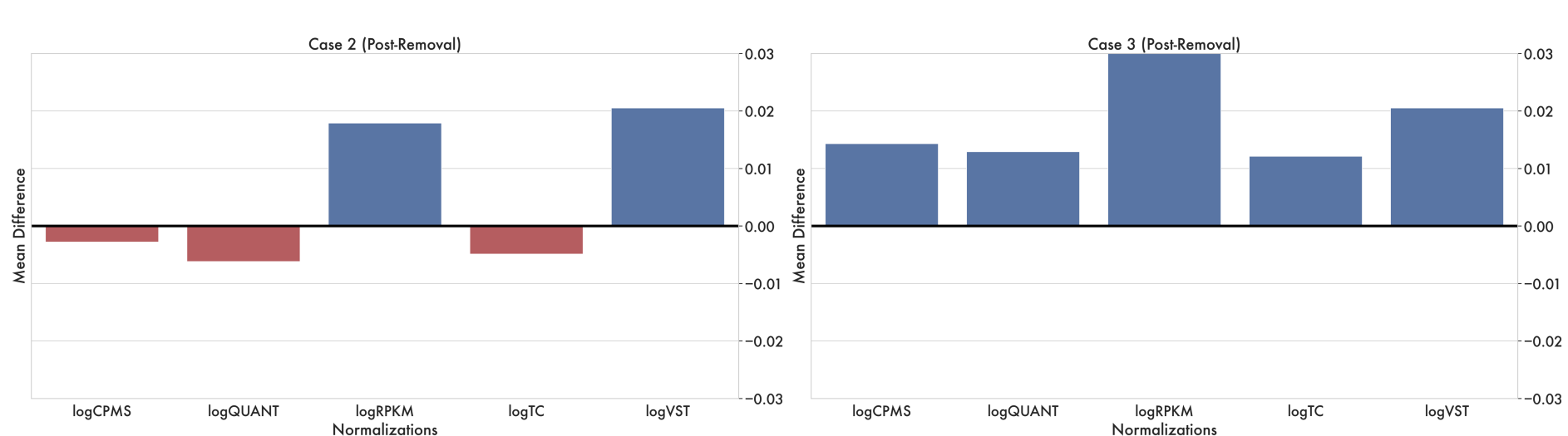
Tissue of origin masks the true information overlap between gene expression and metabolism

Only ~40% of metabolic variance is associated with gene expression

External covariates have significant effect and predictive power with cancer metabolism

Cancer Atlas of Metabolic Profiles

ccRCC3

ccRCC4

Necrosis
- Does Not Exhibit Necrosis
- Exhibits Necrosis

Raw DC

PostpDC

Immune system expression is highly important in cancer metabolism, both *in-vitro* and *in-vivo*

References

1. Hanahan, Douglas, and Robert A. Weinberg. "Hallmarks of Cancer: The Next Generation." Cell, vol. 144, no. 5, Mar. 2011, pp. 646–74, https://doi.org/10.1016/j.cell.2011.02.013.
2. Leone, Robert D., and Jonathan D. Powell. "Fueling the Revolution: Targeting Metabolism to Enhance Immunotherapy." Cancer Immunology Research, vol. 9, no. 3, Mar. 2021, pp. 255–60, https://doi.org/10.1158/2326-6066.cir-20-0791. Accessed 7 Feb. 2022.
3. Li, Huaping, et al. "Gene Expression is a Poor Predictor of Steady-State Metabolite Abundance in Cancer Cells." The FASEB Journal, vol. 36, no. 5, Wiley, Apr. 2022, https://doi.org/10.1096/fj.202101921rr. Accessed 4 Nov. 2023.
4. Ghandi, Mahmoud, et al. "Next-Generation Characterization of the Cancer Cell Line Encyclopedia." Nature, vol. 569, no. 7757, May 2019, pp. 503–8, https://doi.org/10.1038/s41586-019-1186-3. Accessed 20 Jan. 2021.
5. Benedetti, Elisa, et al. "A Multimodal Atlas of Tumour Metabolism Reveals the Architecture of Gene–Metabolite Covariation." Nature Metabolism, vol. 5, no. 6, June 2023, pp. 1029–44, https://doi.org/10.1038/s42255-023-00817-8. Accessed 29 June 2023.
6. Pedregosa, Fabian, et al. "Scikit-Learn: Machine Learning in Python." Journal of Machine Learning, vol. 12, Oct. 2011, pp. 2825–30.
7. Seabold, Skipper, and Josef Perktold. "Statsmodels: Econometric and Statistical Modeling with Python." Proceedings of the 9th Python in Science Conference, 2010.
8. Székely, Gábor J., et al. "Measuring and Testing Dependence by Correlation of Distances." The Annals of Statistics, vol. 35, no. 6, Dec. 2007, pp. 2769–94, https://doi.org/10.1214/009053607000000505.
9. Kornberg, Hans. "Metabolism." Encyclopedia Britannica, 15 Nov. 2018, www.britannica.com/science/metabolism.
[*] Created with Biorender.com